Topics in Computational Inference

David M. Allen
University of Kentucky

October 26, 2018

10 The Logistic Function

This chapter presents the logistic function, a model that seems appropriate many biological phenomenon. It is derived and illustrated here. See De Sapio [1] for an alternate derivation and examples.

Section 10.1 — The Genesis

The genesis of the *logistic function* is a differential equation. Let t be the independent variable and $x(t)$ be the logistic function. The defining equation is

$$\dot{x}(t) = \theta(x(t) - a)(b - x(t)) \tag{10.1.1}$$

where θ, a, and b are constants, and $\dot{x}(t)$ is the *derivative* of $x(t)$ with respect to t. The constant θ characterizes the rate of growth; a and b respectively represent the lower and upper bounds of $x(t)$. The derivative gives the slope of $x(t)$ at any point t. The derivative is also interpreted as the *rate of change* of $x(t)$.

Interpretation

Equation (10.1.1) says:

1. the slope of $x(t)$ approaches zero as $x(t)$ approaches a or when $x(t)$ approaches b;
2. the slope of $x(t)$ has the same sign as θ when $a < x(t) < b$; and
3. the magnitude of the slope of $x(t)$ is larger when $|\theta|$ is larger.

These properties make the logistic function a good candidate for modeling phenomenon where the domain of the independent variable is unbounded and the range of the response is bounded between two values. The logistic function with $a = 0$ and $b = 1$ is frequently used to model probabilities.

Solving for the Logistic Function

The first step for solving the differential equation is manipulation of the equation to facilitate finding the antiderivatives:

$$
\begin{aligned}
\dot{x}(t) &= \theta\,(x(t)-a)(b-x(t)) \\
\frac{1}{\dot{x}(t)} &= \frac{1}{\theta\,(x(t)-a)(b-x(t))} \\
\frac{\theta(b-a)}{\dot{x}(t)} &= \frac{1}{x(t)-a} + \frac{1}{b-x(t)} \\
c &= \frac{\dot{x}(t)}{x(t)-a} + \frac{\dot{x}(t)}{b-x(t)}.
\end{aligned}
$$

The symbol c is substituted for $\theta(b-a)$.

Solution Continued

The second step is the integration and manipulation to get $x(t)$ on the left hand side:

$$d + c\,t = \log(x(t) - a) - \log(b - x(t))$$

$$d + c\,t = \log\left(\frac{x(t) - a}{b - x(t)}\right)$$

$$\exp(d + c\,t) = \frac{x(t) - a}{b - x(t)}$$

The last expression is solved for $x(t)$. The constant d is a constant of integration.

A Special Case

With $a = 0$ and $b = 1$, the solution is particularly simple:

$$x(t) = \frac{\exp(d + c\,t)}{1 + \exp(d + c\,t)}$$

The Logit Function

An intermediate expression in the derivation of $x(t)$ is

$$d + c\,t = \log\left(\frac{x(t) - a}{b - x(t)}\right). \tag{10.1.2}$$

This demonstrates that a function of $x(t)$ is a linear function of t. When $a = 0$ and $b = 1$ the expression on the right hand side is called the *logit function*.

Examples of Logistic Curves

The graphs in Figure 10.1 and Figure 10.2 are to illustrate how changing values of c and d affect the shape of the logistic function. The interpretation of a and b as upper and lower bounds is clear, and the examples fix $a = 0$ and $b = 1$.

The Effect of c

The curves in Figure 10.1 have different values of c. Note that the curve with the larger value of c has the larger slope near the center of the graph. The respective values of d are adjusted so curves cross near the center of the graph.

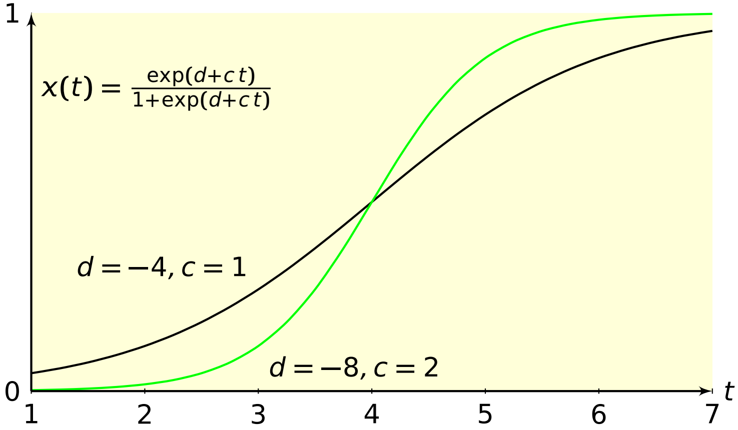


Figure 10.1: Two logistic curves with different c

The Effect of d

The curves in Figure 10.2 have the same value of c, but different values of d. Note that the horizontal distance between the curves is constant.

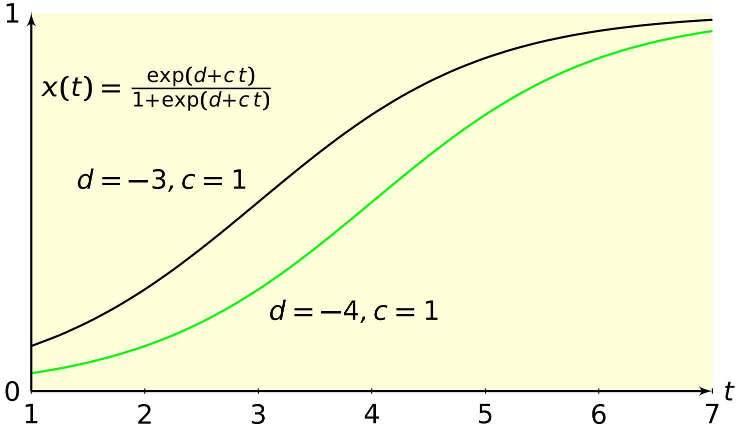


Figure 10.2: Logistic curves with different d

Inverse Estimation

Suppose the constants a, b, c, and d have been estimated from data using a non-linear models program. Suppose further that interest turns to estimating the value of the independent variable, t, that produces a specified response. This is accomplished by replacing $x(t)$ in equation (10.1.2) by the specified response and solving for t.

If the response is specified in terms a weighted average of a and b, the formula for the corresponding t is simplified. Suppose $x(t) = p\,a + q\,b$ where $p + q = 1$ and $0 < p < 1$. Equation (10.1.2) reduces to $d + c\,t = \log(q/p)$.

Section 10.2 — An example

This is an example from Venables and Ripley [2, page 218]. This is an experiment on the toxicity to the tobacco budworm *Heliothis virescens* of doses of pyrethroid *trans*-cypermethrin to which the moths were beginning to show resistance. Batches of 20 moths of each sex were exposed for three days to the pyrethroid and the number in each batch that were dead or knocked down was recorded.

The Data

The data are

Sex	Dose μg					
	1	2	4	8	16	32
Male	1	4	9	13	18	20
Female	0	2	6	10	12	16

Entries in the table are the number out of 20 that were dead or knocked down. Logistic regression is done with the independent variable $\log_2$ of dose.

R Code for Analysis

R code for an analysis is

```
ldose  <- c(0:5, 0:5)
sex    <- c(rep("Male",6),rep("Female",6))
killed <- c(1,4,9,13,18,20,0,2,6,10,12,16)/20

o <- glm(killed ~ factor(sex) + factor(sex):ldose -1,
        family=binomial(link = "logit"),
        weights=rep(20,length(killed)))
print(summary(o))
```

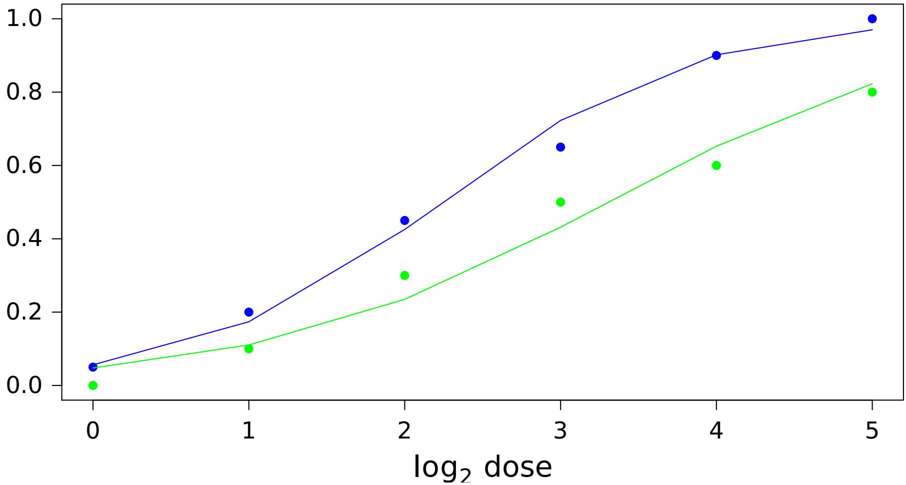
R Code for Graph

```
require(tikzDevice)
# width and height below nearly fill a beamer screen
tikz("budworm-graph.tex",standAlone=FALSE, width=4.8, height=2.5)
par(mex=0.6,mar=c(4.2,4,0,0)+0.1,las=1,cex.axis=0.8, pch=20)

fit <- fitted(o)
plot(c(0,5),c(0,1), type="n", ylab="", xlab="$\\log_2$ dose")
points(ldose[sex=="Male"],   killed[sex=="Male"],   col = "blue")
points(ldose[sex=="Female"], killed[sex=="Female"], col = "green"
lines(ldose[sex=="Male"],    fit[sex=="Male"],      col = "blue")
lines(ldose[sex=="Female"],  fit[sex=="Female"],    col = "green"

dev.off()
```

Fitted Response Over the Data



An Exercise

Exercise 10.1. We have five groups of 25 chickens. Each chicken in each group was administered a measured dose of a drug as shown in the schedule below. Each chicken that fell asleep within 30 minutes of drug administration is declared a responder.

Group	1	2	3	4	5
Dose (*mg*)	10	20	30	40	50
Responders	0	1	16	23	24

Establish the *best* logistic relationship between the proportion of chickens that went to sleep and the dose of the drug.

References

[1] Rodolfo De Sapio. *Calculus for the Life Sciences*. San Francisco: W. H. Freeman and Company, 1978.

[2] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S-PLUS*. Third. The fourth edition has been published. Springer-Verlag, 1999.